

6/1/01



#5

ENTERED

OIPE

RAW SEQUENCE LISTING

DATE: 04/02/2002 8.5

PATENT APPLICATION: US/10/004,357

TIME: 14:03:58

Input Set : A:\02-107010us SEQUENCE LISTING.txt

Output Set: N:\CRF3\04022002\J004357.raw

```

4 <110> APPLICANT: Castle, Linda A.
5   Siehl, Dan
6   Giver, Lorraine
7   Minshull, Jeremy
8   Ivy, Christina
9   Chen, Yong Hong
10  Duck, Nicholas B.
12 <120> TITLE OF INVENTION: NOVEL GLYPHOSATE N-ACETYLTRANSFERASE
13   (GAT) GENES
15 <130> FILE REFERENCE: 02-107010US
17 <140> CURRENT APPLICATION NUMBER: US 10/004,357
18 <141> CURRENT FILING DATE: 2001-10-29
20 <150> PRIOR APPLICATION NUMBER: US 60/244,385
21 <151> PRIOR FILING DATE: 2000-10-30
23 <160> NUMBER OF SEQ ID NOS: 515
25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 441
29 <212> TYPE: DNA
30 <213> ORGANISM: Bacillus licheniformis
32 <400> SEQUENCE: 1
33 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
34 cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcgctttt 120
35 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
36 cattcagagc ttgaaggcga agaacagtat cagctgagag ggatggcgac gcttgaaggga 240
37 taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
38 aagggggcag accttttatg gtgcaatgcc aggacatctg tgagcgggta ctatgaaaag 360
39 ctcggttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
40 atgtataaga aattgacgta a                                     441
42 <210> SEQ ID NO: 2
43 <211> LENGTH: 441
44 <212> TYPE: DNA
45 <213> ORGANISM: Unknown
47 <220> FEATURE:
48 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
50 <400> SEQUENCE: 2
51 atgattgaag tcaaacctat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
52 cggccgaatc agccgcttga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
53 cacctcggcg gatattatcg ggacaggctg atcagcatcg cctcctttca tcaagccgaa 180
54 cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
55 taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
56 aaaggcgagc accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
57 ctcggttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420

```

RAW SEQUENCE LISTING

DATE: 04/02/2002

PATENT APPLICATION: US/10/004,357

TIME: 14:03:58

Input Set : A:\02-107010us SEQUENCE LISTING.txt

Output Set: N:\CRF3\04022002\J004357.raw

```

58 atgtataaga aattgacata a 441
60 <210> SEQ ID NO: 3
61 <211> LENGTH: 441
62 <212> TYPE: DNA
63 <213> ORGANISM: Unknown
65 <220> FEATURE:
66 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
68 <400> SEQUENCE: 3
69 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
70 cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
71 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
72 cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
73 taccgcgagc aaaaagcggg aagtaagctt atccgccatg ccgaagagct tcttcgaaaa 300
74 aaaggcgagg accttttatg gtgcaacgcc aggatatctg tgagcggcta ctatgaaaag 360
75 ctcggcctca gcgaacaagg cgggacttac gacataccgc cgatcggacc tcatattttg 420
76 atgtataaga aattggcata a 441
78 <210> SEQ ID NO: 4
79 <211> LENGTH: 441
80 <212> TYPE: DNA
81 <213> ORGANISM: Unknown
83 <220> FEATURE:
84 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
86 <400> SEQUENCE: 4
87 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
88 cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
89 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
90 cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
91 taccgcgagc aaaaagcggg aagtaagctt atccgccatg ccgaagagct tcttcgaaaa 300
92 aaaggcgagg accttttatg gtgcaacgcc aggatatctg tgagcggcta ctatgaaaag 360
93 ctcggcctca gcgaacaagg cgggacttac gacataccgc cgatcggacc tcatattttg 420
94 atgtataaga aattggcata a 441
96 <210> SEQ ID NO: 5
97 <211> LENGTH: 441
98 <212> TYPE: DNA
99 <213> ORGANISM: Unknown
101 <220> FEATURE:
102 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
104 <400> SEQUENCE: 5
105 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
106 cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
107 cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
108 cattcagagc ttgagggcga agaacagtat cagctgagag ggatggcgac gcttgaaggg 240
109 taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcgaaaa 300
110 aagggggcag accttttatg gtgcaatgcc aggatatctg tgagcggcta ctatgaaaag 360
111 ctcggcctca gcgaacaagg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
112 atgtataaga aattgacgta a 441
114 <210> SEQ ID NO: 6
115 <211> LENGTH: 146
116 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/004,357

DATE: 04/02/2002

TIME: 14:03:58

Input Set : A:\02-107010us SEQUENCE LISTING.txt

Output Set: N:\CRF3\04022002\J004357.raw

```

117 <213> ORGANISM: Bacillus licheniformis
119 <400> SEQUENCE: 6
120 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
121 1 5 10 15
122 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
123 20 25 30
124 Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
125 35 40 45
126 Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
127 50 55 60
128 Glu Gly Glu Glu Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
129 65 70 75 80
130 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
131 85 90 95
132 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
133 100 105 110
134 Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
135 115 120 125
136 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
137 130 135 140
138 Leu Thr
139 145
142 <210> SEQ ID NO: 7
143 <211> LENGTH: 146
144 <212> TYPE: PRT
145 <213> ORGANISM: Unknown
147 <220> FEATURE:
148 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
150 <400> SEQUENCE: 7
151 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
152 1 5 10 15
153 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
154 20 25 30
155 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Asp
156 35 40 45
157 Arg Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
158 50 55 60
159 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
160 65 70 75 80
161 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
162 85 90 95
163 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
164 100 105 110
165 Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
166 115 120 125
167 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
168 130 135 140
169 Leu Thr
170 145

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/004,357

DATE: 04/02/2002

TIME: 14:03:58

Input Set : A:\02-107010us SEQUENCE LISTING.txt

Output Set: N:\CRF3\04022002\J004357.raw

```

173 <210> SEQ ID NO: 8
174 <211> LENGTH: 146
175 <212> TYPE: PRT
176 <213> ORGANISM: Unknown
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
181 <400> SEQUENCE: 8
182 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
183   1           5           10           15
184 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
185           20           25           30
186 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
187           35           40           45
188 Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu
189           50           55           60
190 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
191 65           70           75           80
192 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
193           85           90           95
194 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
195           100          105          110
196 Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Gly
197           115          120          125
198 Ile Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
199           130          135          140
200 Leu Ala
201 145
204 <210> SEQ ID NO: 9
205 <211> LENGTH: 146
206 <212> TYPE: PRT
207 <213> ORGANISM: Unknown
209 <220> FEATURE:
210 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
212 <400> SEQUENCE: 9
213 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
214   1           5           10           15
215 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
216           20           25           30
217 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
218           35           40           45
219 Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu
220           50           55           60
221 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
222 65           70           75           80
223 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
224           85           90           95
225 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
226           100          105          110
227 Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Leu Ser Glu Gln Gly Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/004,357

DATE: 04/02/2002

TIME: 14:03:58

Input Set : A:\02-107010us SEQUENCE LISTING.txt

Output Set: N:\CRF3\04022002\J004357.raw

```

228          115          120          125
229 Ile Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
230          130          135          140
231 Leu Ala
232 145
235 <210> SEQ ID NO: 10
236 <211> LENGTH: 146
237 <212> TYPE: PRT
238 <213> ORGANISM: Unknown
240 <220> FEATURE:
241 <223> OTHER INFORMATION: Unidentified microorgansim derived from soil sample
243 <400> SEQUENCE: 10
244 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
245 1 5 10 15
246 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
247 20 25 30
248 Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly
249 35 40 45
250 Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
251 50 55 60
252 Glu Gly Glu Glu Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
253 65 70 75 80
254 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
255 85 90 95
256 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
257 100 105 110
258 Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
259 115 120 125
260 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
261 130 135 140
262 Leu Thr
263 145
266 <210> SEQ ID NO: 11
267 <211> LENGTH: 441
268 <212> TYPE: DNA
269 <213> ORGANISM: Artificial Sequence
271 <220> FEATURE:
272 <223> OTHER INFORMATION: Synthetic DNA Sequence
274 <400> SEQUENCE: 11
275 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
276 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggtagcttt 120
277 caccctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
278 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
279 taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
280 aaaggcgcgg accttttatg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360
281 ctcggttca gcgaacaggg cgaagtctac gacataaccg cggtcggacc tcatattttg 420
282 atgtataaga aattgacgta a 441
284 <210> SEQ ID NO: 12
285 <211> LENGTH: 441

```

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/004,357

DATE: 04/02/2002

TIME: 14:03:59

Input Set : A:\02-107010us SEQUENCE LISTING.txt

Output Set: N:\CRF3\04022002\J004357.raw

L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:3687 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:200
L:3690 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:200
L:4463 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:243
L:4464 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:243
L:8673 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:387
L:10698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:452
L:10706 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:452
L:12039 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:495
L:12041 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:495
L:12043 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:495